



1

SEQUENCE LISTING

10
<110> Tickle, Ian J
Vonrhein, Clemens
Williams, Pamela A
Jhoti, Harren
Kirton, Stewart Brian

<120> Crystal structure of cytochrome P450

<130> 620-282

<140> US 10/690,991
<141> 2003-10-23

<150> US 60/421,063
<151> 2002-10-25

<150> PCT/GB02/02668
<151> 2002-05-30

<150> US 10/221,036
<151> 2002-04-02

<150> US 60/479,448
<151> 2003-06-19

<160> 6

<170> PatentIn version 3.1

<210> 1
<211> 1458
<212> DNA
<213> Artificial sequence

<220>
<223> Codes for SEQ ID NO: 2, a human 3A4 P450 protein truncated in its N-terminal region to delete the hydrophobic trans-membrane domain, and the region replaced by a short N-terminal sequence.

<400> 1	atggcatacg gtactcattc acatggctcg tttaaaaaac tgggaattcc agggcccaca	60
	cctctgcctt ttttggaaaa tattttgtcc taccataagg gctttgtat gtttgacatg	120
	aatgtcata aaaagtatgg aaaagtgtgg ggctttatg atggtaaca gcctgtgctg	180
	gctatcacag atcctgacat gatcaaaaca gtgctagtga aagaatgtta ttctgtcttc	240
	acaaaccgga ggcctttgg tccagtggaa tttatgaaaa gtgccatctc tatacgtag	300
	gatgaagaat ggaagagatt acgatcattt ctgtctccaa cttcaccag tgaaaaactc	360
	aaggagatgg tccctatcat tgcccagtat ggagatgtgt tggtgagaaa tctgaggcgg	420
	gaagcagaga caggcaagcc tgcacccctt aaagacgtct ttggggccta cagcatggat	480
	gtgatcacta gcacatcatt tggagtgaac atcgactctc tcaacaatcc acaagacccc	540
	tttggaaaa acaccaagaa gcttttaaga tttgattttt tggatccatt ctgtctctca	600
	ataacagtct ttccatttcct catcccaatt cttgaagtat taaatatctg tggtttcca	660
	agagaagtta caaattttt aagaaaatct gtaaaaagga tgaaagaaag tcgcctcgaa	720
	gatacacaaa agcaccgagt ggatttcctt cagctgatga ttgactctca gaattcaaaa	780

gaaactgagt	cccacaaagc	tctgtccgat	ctggagctcg	tggcccaatc	aattatcttt	840
attttgctg	gctatgaaac	cacgagcagt	gttctctcct	tcattatgta	tgaactggcc	900
actcaccctg	atgtccagca	gaaactgcag	gaggaaattg	atgcagttt	acccaataag	960
gcaccaccca	cctatgatac	tgtgctacag	atggagtatc	ttgacatggt	ggtgaatgaa	1020
acgctcagat	tattccaaat	tgctatgaga	cttgagaggg	tctgaaaaaa	agatgtttag	1080
atcaatggga	tgttcattcc	caaaggggtg	gtggtgatga	ttccaagcta	tgctcttcac	1140
cgtgacccaa	agtactggac	agagcctgag	aagttcctcc	ctgaaagatt	cagcaagaag	1200
aacaaggaca	acatagatcc	ttacatatac	acacccttg	gaagtggacc	cagaaactgc	1260
atggcatga	ggttgctct	catgaacatg	aaacttgctc	taatcagagt	ccttcagaac	1320
tttccttca	aaccttgtaa	agaaacacag	atccccctga	aattaagctt	aggaggactt	1380
cttcaaccag	aaaaaccgt	tgttctaaag	gttgagtcaa	gggatggcac	cgtaaagtgg	1440
gccaccatc	accattga					1458

<210> 2
 <211> 485
 <212> PRT
 <213> Artificial sequence

<220>
 <223> Human 3A4 P450 protein truncated in its N-terminal region to
 delete the hydrophobic trans-membrane domain, and the region
 replaced by a short N-terminal sequence.

<400> 2

Met Ala Tyr Gly Thr His Ser His Gly Leu Phe Lys Lys Leu Gly Ile
 1 5 10 15

Pro Gly Pro Thr Pro Leu Pro Phe Leu Gly Asn Ile Leu Ser Tyr His
 20 25 30

Lys Gly Phe Cys Met Phe Asp Met Glu Cys His Lys Lys Tyr Gly Lys
 35 40 45

Val Trp Gly Phe Tyr Asp Gly Gln Gln Pro Val Leu Ala Ile Thr Asp
 50 55 60

Pro Asp Met Ile Lys Thr Val Leu Val Lys Glu Cys Tyr Ser Val Phe
 65 70 75 80

Thr Asn Arg Arg Pro Phe Gly Pro Val Gly Phe Met Lys Ser Ala Ile
 85 90 95

Ser Ile Ala Glu Asp Glu Glu Trp Lys Arg Leu Arg Ser Leu Leu Ser
 100 105 110

Pro Thr Phe Thr Ser Gly Lys Leu Lys Glu Met Val Pro Ile Ile Ala
 115 120 125

Gln Tyr Gly Asp Val Leu Val Arg Asn Leu Arg Arg Glu Ala Glu Thr
130 135 140

Gly Lys Pro Val Thr Leu Lys Asp Val Phe Gly Ala Tyr Ser Met Asp
145 150 155 160

Val Ile Thr Ser Thr Ser Phe Gly Val Asn Ile Asp Ser Leu Asn Asn
165 170 175

Pro Gln Asp Pro Phe Val Glu Asn Thr Lys Lys Leu Leu Arg Phe Asp
180 185 190

Phe Leu Asp Pro Phe Phe Leu Ser Ile Thr Val Phe Pro Phe Leu Ile
195 200 205

Pro Ile Leu Glu Val Leu Asn Ile Cys Val Phe Pro Arg Glu Val Thr
210 215 220

Asn Phe Leu Arg Lys Ser Val Lys Arg Met Lys Glu Ser Arg Leu Glu
225 230 235 240

Asp Thr Gln Lys His Arg Val Asp Phe Leu Gln Leu Met Ile Asp Ser
245 250 255

Gln Asn Ser Lys Glu Thr Glu Ser His Lys Ala Leu Ser Asp Leu Glu
260 265 270

Leu Val Ala Gln Ser Ile Ile Phe Ile Phe Ala Gly Tyr Glu Thr Thr
275 280 285

Ser Ser Val Leu Ser Phe Ile Met Tyr Glu Leu Ala Thr His Pro Asp
290 295 300

Val Gln Gln Lys Leu Gln Glu Glu Ile Asp Ala Val Leu Pro Asn Lys
305 310 315 320

Ala Pro Pro Thr Tyr Asp Thr Val Leu Gln Met Glu Tyr Leu Asp Met
325 330 335

Val Val Asn Glu Thr Leu Arg Leu Phe Pro Ile Ala Met Arg Leu Glu
340 345 350

Arg Val Cys Lys Lys Asp Val Glu Ile Asn Gly Met Phe Ile Pro Lys
355 360 365

Gly Val Val Val Met Ile Pro Ser Tyr Ala Leu His Arg Asp Pro Lys
370 375 380

Tyr Trp Thr Glu Pro Glu Lys Phe Leu Pro Glu Arg Phe Ser Lys Lys
385 390 395 400

Asn Lys Asp Asn Ile Asp Pro Tyr Ile Tyr Thr Pro Phe Gly Ser Gly
 405 410 415

Pro Arg Asn Cys Ile Gly Met Arg Phe Ala Leu Met Asn Met Lys Leu
 420 425 430

Ala Leu Ile Arg Val Leu Gln Asn Phe Ser Phe Lys Pro Cys Lys Glu
 435 440 445

Thr Gln Ile Pro Leu Lys Leu Ser Leu Gly Gly Leu Leu Gln Pro Glu
 450 455 460

Lys Pro Val Val Leu Lys Val Glu Ser Arg Asp Gly Thr Val Ser Gly
 465 470 475 480

Ala His His His His
 485

<210> 3
 <211> 16
 <212> PRT
 <213> Artificial sequence

<220>
 <223> N-terminal sequence for expression of human 3A4 P450

<400> 3

Met Ala Tyr Gly Thr His Ser His Gly Leu Phe Lys Lys Leu Gly Ile
 1 5 10 15

<210> 4
 <211> 33
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 4
 ggaattcata tggctctcat cccagacttg gcc 33

<210> 5
 <211> 47
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer

<400> 5
 tgcggtcgac tcaatggta tggggctc cacttacggt gccatcc 47

<210> 6
<211> 69
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 6
ttaacatatg gcatatggta otcattcaca tggtctgttt aaaaaactgg gaattccagg 60
gccccacacc 69